DOCUMENT RESUME

ED 339 731 TM 017 635

AUTHOR Cashin, William E.
TITLE Defining and Evaluating College Teaching. Idea Paper No. 21.
INSTITUTION Kansas State Univ., Manhattan. Center for Faculty Evaluation and Development in Higher Education.
PUB DATE Sep 89
NOTE 6p.
AVAILABLE FROM Center for Faculty Evaluation and Development, Kansas State University, Manhattan, KS 66505 (individual copies for $1.00, a complete set of Idea Papers for $5.00, 25-99 copies $0.25 per copy, and 100 or more at $0.20 per copy. Orders of less than $25.00 must be prepaid).
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *College Instruction; *Definitions; Educational Assessment; Evaluation Methods; Higher Education; *Instructional Effectiveness; *Professors; Student Evaluation of Teacher Performance; Teacher Effectiveness; *Teacher Evaluation; Teaching Methods

ABSTRACT

A review of data that have been used to evaluate college teaching suggests that the definition of teaching used for those evaluations has been very incomplete. An expanded definition includes the following seven components of college teaching: (1) subject matter mastery; (2) curriculum development; (3) course design; (4) delivery of instruction; (5) assessment of instruction; (6) availability of teacher to students; and (7) administrative requirements. Several possible sources of data exist to evaluate college teaching. Data from students have generally been collected systematically, but such data usually focus on instruction. Sources include self-ratings by teachers, personnel files, student evaluations, peers (same subject matter), colleagues (other subject matter), ratings by department chair or dean, administrator ratings, evaluation by instructional consultants, and other sources unique to the given aspect of teaching. It is evident that evaluating college teaching adequately requires more time than has generally been given in the past. An 11-item list of references is included. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

WILLIAM E. CASHIN

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED339731

DEFINING AND EVALUATING COLLEGE TEACHING

IDEA Paper No. 21

William E. Cashin
Kansas State University

TM017635

BEST COPY AVAILABLE

IDEA Paper No. 21

center for FACULTY EVALUATION & DEVELOPMENT
Division of Continuing Education
Kansas State University

September, 1989

Defining and Evaluating College Teaching

William E. Cashin
Kansas State University

Almost every institution of higher education states somewhere that the primary purpose of the institution is **teaching**. While commitment to teaching quality certainly varies across institutions, every college or university does evaluate teaching in some way. More and more, higher education's various publics (students, parents, legislators, and others) are insisting that we pay more than lip service to this commitment, that teaching be evaluated seriously and substantively. The time has come for higher education to put its actions where its rhetoric is. I believe that the vast majority of colleges and universities are willing to do this. The problem is: **how**?

This paper begins by reviewing the various kinds of data that have been used to evaluate college teaching both for personnel decisions and for improvement. An examination of those data suggests that the evaluation of college teaching has been based on a **very incomplete** definition of teaching. I will suggest a more comprehensive definition—or description—of college-level teaching, and list seven areas to be considered when evaluating teaching. Several aspects or facets will be listed for each area. Possible sources—mainly people— will also be suggested, sources that might provide a basis to evaluate each area.

Previous Research

In 1976 Centra (1977, see also 1979) surveyed 453 department heads. He asked them to rate the current use and importance of fifteen kinds of data in evaluating teaching. Centra's survey not only provides a useful list of possible sources of data, it also summarizes most of the kinds of data used in both previous and subsequent research. For example, as far back as 1961, Gustad (1967) used essentially the same list. Seldin continued to use the list in his 1978 and 1983 surveys of liberal arts colleges (see Seldin, 1980 and 1984). Below are Centra's fifteen kinds of data, listed in their order of importance as rated by all 453 department heads.

1st	Chairman evaluation
2nd (tied)	Colleagues' opinions
2nd (tied)	Systematic student ratings
4th	Committee evaluation
5th	Informal student opinions
6th	Dean evaluation
7th	Content of course syllabi and examinations
8th	Popularity of elective courses (for example, enrollment)
9th	Self-evaluation or report
10th	Teaching improvement activities
11th	Student examination performance
12th	Colleague ratings based on classroom visits
13th	Alumni opinions or ratings
14th	Long-term follow up of how students perform
15th	Videotape of classroom teaching.

Only the first six of these sources were rated on average as being of **major importance**. Everything after "dean evaluation" was considered to be of **minor importance**.

Centra's (1977) list has the obvious advantage of outlining the kinds of data that have been used by colleges and universities to evaluate teaching. A second reading, however, may give rise to some reservations. It seems to me that at least four of them: chairman evaluation, colleagues' opinions, committee evaluation, and dean evaluation are **not data** but **evaluators**. How do the chairman and others know whether a faculty member is an effective teacher? Do they depend upon the complaints of a few students, gossip in the faculty lounge, or comments overheard through a open classroom door? Or do they use systematic student ratings, surveys of the department faculty, and classroom observations. I suspect that most of the data actually used by the chairperson, colleagues, committees, and the dean are somewhere else on the list.

Others ask the question: is the list complete? Arreola (1986, 1989) argues— persuasively I think—that a complete definition of teaching should include three broad dimensions: 1) content expertise, 2) instructional delivery skills and characteristics, and 3) instructional design skills. Most of the data on Centra's (1977) list are mainly concerned with Arreola's second dimension: different aspects of classroom instruction.

An Expanded Definition of Teaching

Building upon Arreola's (1986, 1989) and Centra's (1977) work, I suggest that college teaching involves the **seven different areas** (listed below in roughly chronological order). Under each area are listed some aspects or facets of that area.

1. Subject matter mastery
   a. content areas
   b. comprehensiveness of content
   c. currency of content
   d. objectivity of coverage

2. Curriculum development
   a. courses' fit with other courses
   b. course revisions
   c. new courses developed

3. Course design
   a. instructional goals and objectives
   b. content coverage
   c. appropriateness of teaching methods
   d. appropriateness of assessment methods

4. Delivery of instruction
   a. method(s), e.g., lecture, PSI, labs
   b. skills, e.g., speaking, explaining
   c. aids, e.g., handouts, AV

Copyright 1989 Center for Faculty Evaluation and Development

5. Assessment of instruction
   a. tests, e.g., multiple-choice, essay, oral
   b. papers, projects
   c. practicums
   d. grading practices

6. Availability to students
   a. office hours
   b. other, e.g., telephone hours before tests
   c. informal, e.g., meets students for coffee, meals

7. Administrative requirements (completed and on time)
   a. book orders
   b. library reserve
   c. syllabi on file
   d. comes to class
   e. grade reports (probation, final, etc.)

If the above is an appropriate list of the various aspects of college teaching, I would contend that U. S. higher education has a serious problem. Many believe—and I share that belief—that student ratings are the only **primary data** that are **systematically** gathered at many colleges and universities. Most student rating forms are mainly concerned with the delivery of instruction. Some forms also provide information on the assessment of instruction, and occasionally they cover the instructor's contacts with students outside of class. Nevertheless, typically they only cover three of the seven aspects or areas.

Data to Evaluate College Teaching

Table 1 is an attempt to organize possible sources of data to evaluate college teaching. Seven areas of teachings and their various aspects or facets are listed along the vertical dimension, and seven sources of data—mostly people who can supply or judge data—are listed along the horizontal dimension. In the intersecting cells are examples of possible kinds of data— indicated by letters explained at the bottom of the table. The letters refer to specific kinds of data relevant to the area and aspect of teaching and to the supplied or judged by a given group of people. Where a cell is left blank, the suggestion is that the source cannot provide data. (Where there is a "?," I am suggesting that the source may sometimes provide data but not usually.)

Self. The instructor **himself** or **herself** is listed first because he or she has knowledge about **all** of the areas of teaching, and there are some things that only the instructor may know. Typically the instructor's information is obtained by means of a **self-report** containing the instructor's comments about the various areas. The self-report should be primarily **descriptive** rather than **evaluative**. We are not so much interested in whether the instructor thinks the teaching went well or poorly as to **why** it went well or poorly. The instructor's self-report should be compared with the other data. Often a self-report can provide a perspective that none of the other data can.

Files. This general term includes all of the data that can be kept in the institution's **personnel files**, or in **portfolios** or **dossiers** maintained by individual instructors or by departments. An important advantage of files is that most of these data have already been generated for other purposes and requires little additional effort beyond keeping an extra copy. Relevant file data might include **degrees, certificates, and licenses** earned and course materials. **Course materials** should not be restricted to just the course syllabus, nor even to detailed instructional objectives, but should include handouts, objective exams and computer scored results of those exams, or graded examples of essay tests or papers. Actually, the instructor's teaching notes could be a significant source of data, but the department would have to discuss and agreed upon using them ahead of time.

Students. Three different kinds of data are listed in the students' column. **Student ratings** are the most common form of data obtained from students. These ratings include not just numerical ratings, but the students' responses to open-ended questions like "What suggestions do you have about how this course might be improved?" Because they are so time consuming to analyze properly for personnel decisions, I suggest responses to open-ended questions be used only by the instructor for improvement. **Student interviews**, either during or at the end of the term, are a less frequently used means of collecting data from students. Although faculty are inclined to place more credence in these interviews, they probably do not yield markedly different information from student ratings (Braskamp, Ory, & Pieper, 1981; see also Braskamp, Brandenburg, & Ory, 1984) and so I would not recommend them as a routine way to collect student feedback because it is so time consuming. However, such interviews can be very useful in resolving discrepancies found in the data regularly collected. **Student comments and letters**, and other informal feedback from students are still frequently used sources of data at many, especially smaller, institutions. Typically these depend on "volunteers," i.e., we wait for the students to take the initiative and come in to complain or to write a letter. Even when letters are solicited from a random sample of students, the return rates are low. The major question about such data is: How representative of the class as a whole are these students? When student comments differ from other data, then additional data are needed.

Peers. I would like to restrict the term "peers" to faculty members **knowledgeable in the subject matter**. They might be faculty at the same institution or at another institution. In Table 1 I suggest that the judgment of peers can be a useful source of information in at least five of the seven teaching areas. In every case, **reviews of course material** is one basis for that judgment—remembering everything listed above that should be included under course materials. A well-developed syllabus will include the objectives of the course which in turn will serve as the primary context for the evaluation. **Personal contact** has been listed for peers, and for other individuals in Table 1, because it is a source of data that is often available—and probably all too often the primary data used. In my opinion, data based on personal contact should be used with extreme caution because we often do not know how representative that behavior is of the instructor's general teaching behavior. An obvious exception to this would be where the peer had team taught with the instructor. Having sounded this caution, data obtained from personal contact—especially if it is extensive—can often flesh out the other data and provide a more understandable picture of an instructor's teaching.

Faculty are often heard to complain that those who evaluate their teaching have never seen them teach. I find the logic of their argument persuasive (but wonder why faculty forget it when they start talking about student ratings). **Classroom observation** provides an obvious solution to the complaint, or failing that the use of **video- and audiotapes**. However, there are some serious problems when using these methods to evaluate teaching for personnel decisions, especially concerning 1) what context to use in evaluating the teaching, e.g., teaches like me; 2) variability among judges; and 3) representativeness of the class(es) observed. The recommended solution is to have three or more people observe three or more classes—which becomes very expensive of peoples' time. As with some of the other sources of data already discussed, classroom observation and tapes may serve best when reserved to resolve conflicts raised by other, routinely collected data.

Colleagues. The term "colleagues" is used for all of those faculty—those familiar with higher education's academic enterprise—but not knowledgeable of the specific subject mat-

ter. As I have suggested in Table 1, I think there are a number of aspects of teaching which can be judged by colleagues based on the same kinds of data used by peers.

Chair/dean. This term is used to identify especially that person who is the faculty member's immediate academic supervisor, whether it be an academic chairperson or department head, the director of a division, or the dean of a school or college. It can also include others above the instructor's immediate supervisor. As with colleagues, these supervisors can base their judgment on the same data used by the peers. One additional source of data that is often available to academic supervisors that may not be available to peers and colleagues, namely **Department/ Division/College files**. Such files will often contain information relevant to the faculty member's fulfillment of administrative requirements, e.g., getting paperwork in on time. (Faculty often consider paperwork to be of trivial importance, but neglecting it can result in students having no textbooks or assigned readings, or lacking information needed to make registration decisions. These are not trivial as far as the student's education is concerned!)

Administrators. This category is used for other administrators who may not have any supervisory relationship to the faculty member, but may have useful information about how the faculty member fulfills some of his or her teaching responsibilities. Examples are the manager of the Book Store, the librarians, the Registrar, and I have included the Department secretary—without whom nothing would get administered!

Instructional consultant. The majority of colleges and universities do not have instructional consultants, a faculty member or other professional available to help interested faculty improve their teaching. On those campuses which do have access to

Table 1
Data to Evaluate College Teaching

	Sources								
Areas	Self	Files	Students	Peers	Colleagues	Chair/ Dean	Instruct. Admin.	Consultant	Others
Subject matter mastery									
content areas	a	bc		de		?			
comprehensiveness	a	bc		de		?			
currency	a	bc		de		?			
objectivity	a	c		de		?			
Curriculum development									
fit w/ other courses	a	c		de	de	de		?	
course revisions	a	c		de		de			f
new courses	a	c		de		de			f
Course design									
instructional goals	a	c	?	deg	?	deg		dg	
content coverage	a	c	?	deg	?	?		dg	
teaching methods	a	c	?	deg	?	deg		dg	
assessment methods	a	c	?	de	?	deg		dg	
Delivery of instruction									
methods	a		hij	deg	deg			dg	
skills	a		hij	deg	deg			dg	
aids	a	c	hij	deg	deg			dg	
Assessment of instruction									
tests	a	c	hij	deg	deg	deg		dg	
papers, projects	a	c	hij	deg	deg	deg		dg	
practicums	a	c	hij	deg	deg	deg		dg	
grading practices	a	c	hij	deg	deg	deg		dg	
Availability to students									
office hours	a	k	hij	?		?			
other	a		hij	?		?			
informal contacts	a		hij	?		?			
Administrative requirements									
book orders	a	l	hij			m	n		
library reserve	a	l	hij			m	o		
syllabi on file	a	l	hij			m	p		m
comes to class	a		hij	?		?			
grade reports	a	l	hij			m	q		

Note. A ? suggests the person(s) may be a source of data in some cases.

Source's of Data
[a]Self-report
[b]Degrees, certificates, licenses, etc.
[c]Course materials on file
[d]Review of course materials
[e]Personal contact
[f]Community Advisory Committee minutes, letters, etc.
[g]Classroom observation, video- or audiotapes
[h]Student ratings
[i]Student interviews
[j]Student comments or letters
[k]Posted office hours
[l]Instructor's dated copies
[m]Department/Division/ College files
[n]Book Store manager
[o]Librarians
[p]Appropriate secretary
[q]Registrar

such a consultant, these consultants can also offer a judgment about an instructor's teaching based on the same kind of data used by the colleagues and academic supervisors. One caveat, there can be a conflict of interest where the consultant is working with an instructor to improve, then is asked to make a judgment to be used for personnel decisions. The conventional wisdom—which I strongly support—is that instructional consultants only offer their judgments to the individual instructor for his or her improvement. If the consultant does supply data for personnel decisions, it is done only at the written request of the instructor.

Others. The seventh group is labeled "others." It has been included to draw our attention to the possibility of including data that may be unique to a given aspect or facet of teaching. One example related to curriculum development is Community Advisory Committee minutes, letters, etc. More and more universities and four-year colleges are doing what community colleges have done for years, they are forming committees comprised in part of members from the community. These committees offer advice about how the institution might more effectively serve the educational needs of the community. Only two examples are listed in this column but there are most likely others.

Conclusion

This paper has considered only some of the issues in evaluating college teaching, primarily identifying the various areas of teaching responsibility. Perhaps the major issue not discussed is criteria: what constitutes effective teaching? One possible answer is: all of those instructor behaviors which help students learn. Many readers may find that answer too general to be helpful. For the individual course I suggest that the primary context of the evaluation be the goals and objectives of the course. To the extent that the teaching has helped the students achieve those objectives, it should be rewarded. For other areas of teaching, e.g., curriculum development, the goals and objectives of the Department, and of the institution as a whole should be the primary context.

Another issue which has not been discussed is standards: what constitutes "satisfactory" performance, or "outstanding" performance? For example, what kinds of student ratings are considered satisfactory? On a 5-point scale, is 3.0 satisfactory even though the average is 3.5? These questions must be answered at least on an institutional basis. Given all of the differences among academic disciplines, I strongly recommend that they be answered on a departmental basis, with the approval of the respective deans and chief academic officers.

Finally, there is the question of the amount of time it will take to evaluate. Some readers may feel that it will take so much time to evaluate that there will be no time left to teach! It is not as bad as it appears. Most teaching aspects will not have to be evaluated completely every year. An obvious example is degrees, etc. earned; these only need to be updated when there is a change. The same applies to aspects like developing a new course. But even data like course materials would not be completely evaluated every year, e.g., an instructor whose course materials were rated satisfactory need only be evaluated every three to five years, unless the instructor requests otherwise, or there has been a significant change in the course or in the students. And any institution deciding to adopt some of the suggestions made in this paper can phase them in over three or more years.

One implication of this paper is that evaluating teaching will take much more effort than has been given to it in the past, but doing so will yield more complete and more accurate evaluations. The hope is that by better defining and assessing teaching, higher education will be able to develop a system to identify and reward effective teachers.

References

Two books, listed below, deal specifically with evaluating teaching: Braskamp, et al. 1984; and Doyle, 1983. Also of special interest are: Centra, 1979; Miller, 1987; and Seldin, 1980 which discuss faculty evaluation in general.

Arreola, R. A. (1986). Evaluating the dimensions of teaching. *Instructional Evaluation, 8*, 4-12.

Arreola, R. A. (1989). Defining and evaluating the elements of teaching. *Proceedings of Academic Chairpersons: Evaluating Faculty, Students, and Programs* (pp. 1-14). Manhattan: Kansas State University.

Braskamp, L. A., Brandenburg, D. C., & Ory, J. C. (1984). *Evaluating teaching effectiveness: A practical guide*. Beverly Hills, CA: Sage.

Braskamp, L. A., Ory, J. C., & Pieper, D. M. (1981). Student written comments: Dimensions of instructional quality. *Journal of Educational Psychology, 73*, 65-70.

Centra, J. A. (1977). *How universities evaluate faculty performance: A survey of department heads* (Report GREB No. 75-5bR). Princeton, NJ: Educational Testing Service.

Centra, J. A. (1979). *Determining faculty effectiveness: Assessing teaching, research, and service for personnel decisions and improvement*. San Francisco: Jossey-Bass.

Doyle, K. O. (1983). *Evaluating teaching*. Lexington, MA: D. C. Heath.

Gustad, J. W. (1967). Evaluation of teaching performance: Issues and possibilities. In C. B. T. Lee (Ed.), *Improving college teaching* (pp. 265-281). Washington, DC: American Council on Education.

Miller, R. I. (1987). *Evaluating faculty for promotion and tenure*. San Francisco, CA: Jossey-Bass.

Seldin, P. (1980). *Successful faculty evaluation programs: A practical guide to improve faculty performance and promotion/tenure decisions*. Crugers, NY: Coventry Press.

Seldin, P. (1984). *Changing practices in faculty evaluation: A critical assessment and recommendations for improvement*. San Francisco: Jossey-Bass.

IDEA Papers may be ordered from the Center. Individual copies are $1.00. A complete set of IDEA Papers may be ordered for $5.00. Bulk orders: 25-99 copies are 25 cents per copy; 100 or more copies, 20 cents per copy. Orders of less than $25.00 must be prepaid.